

CLAIMS

What is claimed is:

1. A multi-processor system comprising:
a requesting node that provides a first request for data to a home node, the requesting node being operative to provide a second request for the data to at least one predicted node in parallel with first request, the requesting node receiving at least one coherent copy of the data from at least one of the home node and the at least one predicted node.
2. The system of claim 1, further comprising an owner predictor programmed to determine the at least one predicted node that is associated with the first request.
3. The system of claim 2, wherein the requesting node further comprises the owner predictor, the owner predictor being operative to determine the at least one predicted node based on the first request being provided as a speculative type of request.
4. The system of claim 1, wherein a cached copy of the data exists in the system, the home node issuing a third request for the data to an owner node having the cached copy of the data.
5. The system of claim 4, wherein the system employs a directory-based cache coherency protocol, the home node further comprising a directory that maintains directory state information associated with the data, the home node issuing the third request to the owner node based on the directory state information indicating that the owner node has an exclusive cached copy of the data.
6. The system of claim 4, wherein the owner node provides a response to one of (i) the home node and (ii) the home node and the requesting node, the owner node providing the response based on a state of the cached copy of the data at the owner node.
7. The system of claim 4, wherein requesting node provides the first request to the home node in a request channel and the second request to the at least one predicted node in a forward channel, the home node issuing the third request to the owner node in the forward

channel, and each of the owner node and the at least one predicted node providing a respective response in a response channel.

8. The system of claim 7, wherein the system employs a directory-based cache coherency protocol in which requests for the data in the forward channel do not block.

9. The system of claim 4, wherein the at least one predicted node comprises the owner node having an exclusive cached copy of the data, the owner node providing a data response to the requesting node based on which of the second request and the third request arrives at the owner node first.

10. The system of claim 9, wherein the owner node provides a victim message to the home node and the data response to the requesting node in response to the second request arriving at the owner node prior to the third request, the home node providing a speculation acknowledgement to the requesting node in response to the victim message from the owner node.

11. The system of claim 9, wherein the owner node provides a victim message to the home node in response to the third request arriving at the owner node prior to the second request, the owner node providing the data response to the requesting node in response to the third request from the home node.

12. The system of claim 4, wherein the at least one predicted node further comprises a target node having a cache that includes the data having one of an invalid state and a shared state, the at least one predicted node providing a miss response to the requesting node in response to the second request, and the owner node providing a data response to the requesting node in response to the third request.

13. A multi-processor network comprising:
 - a requesting processor that provides a first request for data to a home node;
 - the home node comprising a directory that maintains directory state information associated with the data, the home node being operative to provide a second request to an owner processor in response to the first request if the directory state information indicates that the owner processor has a cached copy of the data; and
 - an owner predictor associated with the requesting processor, the owner predictor being operative to identify at least one predicted node associated with the first request, the first processor providing a third request for the data to the least one predicted node substantially concurrently with the first request.
14. The network of claim 13, wherein the requesting processor is operative to selectively provide the first request as one of a speculative type of request and a non-speculative type of request, the owner predictor being operative to determine the at least one predicted node and enable the requesting processor to provide the third request if the first request is provided as the speculative type of request.
15. The network of claim 13, wherein the owner processor provides a response based on a state of the cached copy of the data at the owner processor, the owner processor providing one of (i) a non-data response to the home node and (ii) data responses to the home node and to the requesting processor.
16. The network of claim 13, wherein the first request is provided in a request channel, and the second and third requests are each provided in a forward channel.
17. The network of claim 16, wherein the data is returned to the requesting processor as a response in a response channel, the response being provided by one of the home node and the owner processor based on whether the owner processor has an exclusive cached copy of the data.
18. The network of claim 16, wherein the system employs a directory-based cache coherency protocol in which requests in the forward channel do not block.

19. The network of claim 13, wherein the at least one predicted node comprises the owner processor having an exclusive cached copy of the data, the owner processor providing a data response to the requesting processor in response to which of the second request and the third request that arrives at the owner node first.

20. The network of claim 19, wherein the owner processor provides a victim message to the home node and the data response to the requesting node in response to the third request arriving at the owner processor prior to the second request, the home node providing a speculation acknowledgement to the requesting processor in response to the victim message from the owner processor.

21. The network of claim 19, wherein the owner processor provides a victim message to the home node in response to the second request arriving at the owner processor prior to the third request, the owner processor also providing the data response to the requesting processor in response to the second request from the home node.

22. The network of claim 13, wherein the at least one predicted node further comprises a target node having a cache that includes the data having one of an invalid state and a shared state, the at least one predicted node providing a miss response to the requesting node in response to the second request, and the owner processor providing a data response to the requesting node in response to the third request.

23. The network of claim 13, further comprising an unordered network interconnect that enables communication of requests and responses among the requesting processor, the at least one predicted node, the owner processor and the home node.

24. A multi-processor system comprising:

means for providing a first request for requested data from a requester to a home node;

means for selectively providing a second request to a predicted target node for the requested data based on the first request, the second request being provided in parallel with the first request; and

means for providing a coherent copy of the requested data to the requester in response to at least one of the first request and the second request.

25. The system of claim 24, further comprising:

means for ascertaining whether an owner node has an exclusive cached copy of the requested data; and

means for providing a third request for the data from the home node to the owner node when the owner node has the exclusive cached copy of the requested data.

26. The system of claim 25, wherein the means for providing the coherent copy of the requested data comprises one of (i) the home node, in response to the first request, when no exclusive cached copy of the requested data exists and (ii) the owner node, in response to one of the second request and the third request, when the owner node has the exclusive cached copy of the requested data.

27. The system of claim 26, further comprising means for providing a response that includes the coherent copy of the requested data from the owner node to the requester and for sending a corresponding victim acknowledgement message from the owner node to the home node in response to the second request.

28. The system of claim 26, further comprising means for changing a state of the cached copy of the requested data at the owner node in response to providing the response that includes the coherent copy of the requested data from the owner node to the requester.

29. A method comprising:
issuing a first request for a block of data from a requester to a home node;
concurrently issuing a second request for the block of data from the requester to a predicted node based on the first request; and
receiving at least one coherent copy of the block of data at the requester from an owner node, if the owner node has an exclusive cached copy of the block of data, and from the home node, if no exclusive cached copy of the block of data exists when the home node receives the first request.
30. The method of claim 29, further comprising issuing a third request for the block data from the home node to the owner node in response to determining that the owner node has the exclusive cached copy of the requested data.
31. The method of claim 30, further comprising providing the coherent copy of the requested data in response to the second request when owner node receives the second request prior to the third request.
32. The method of claim 30, further comprising providing the coherent copy of the requested data in response to the third request when owner node receives the third request prior to the second request.
33. The method of claim 30, further comprising employing a cache coherency protocol wherein the first request is provided in a request channel, the second and third requests are each provided in a forward channel, the coherent copy of the block of data is returned to the requester in a response channel, and wherein requests for the requested data in the forward channel do not block.
34. The method of claim 29, further comprising:
providing a response that includes the coherent copy of the requested data from the owner node to the requester in response to the second request; and
sending a corresponding victim acknowledgement message from the owner node to the home node in response to the second request.

35. The method of claim 29, further comprising:

selectively providing the first request as one of a speculative type of request and a non-speculative type of request; and

the second request being concurrently issued to the at least one predicted node only if the first request is provided as the speculative type of request.